(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) Publication number: 0 674 000 A2

(12) **EUROPEAN PATENT APPLICATION**

(21) Application number: 95104301.7

(22) Date of filing: 23.03.95

(51) Int. Cl.[6]: C12N 9/10, C12N 15/54

(30) Priority: 24.03.94 JP 53804/94
25.11.94 JP 315572/94

(43) Date of publication of application:
27.09.95 Bulletin 95/39

(84) Designated Contracting States:
BE CH DE FR GB IT LI SE

(71) Applicant: TOYOTA JIDOSHA KABUSHIKI KAISHA
1, Toyota-cho
Toyota-shi
Aichi-ken (JP)

(72) Inventor: Nishino, Tokuzo, c/o Toyota Jidosha K. K.
1, Toyota-cho
Toyota-shi,
Aichi (JP)
Inventor: Ohnuma, Shinichi, c/o Toyota Jidosha K. K.
1, Toyota-cho
Toyota-shi,
Aichi (JP)
Inventor: Suzuki, Manabu, c/o Toyota Jidosha K. K.
1, Toyota-cho
Toyota-shi,
Aichi (JP)
Inventor: Ohto, Chikara, c/o Toyota Jidosha K. K.
1, Toyota-cho
Toyota-shi,
Aichi (JP)
Inventor: Asada, Chika, c/o Toyota Jidosha K. K.
1, Toyota-cho
Toyota-shi,
Aichi (JP)
Inventor: Higuchi, Yuka, c/o Toyota Jidosha K. K.
1, Toyota-cho
Toyota-shi,
Aichi (JP)
Inventor: Takeuchi, Yoshie, c/o Toyota Jidosha K. K.
1, Toyota-cho
Toyota-shi,
Aichi (JP)

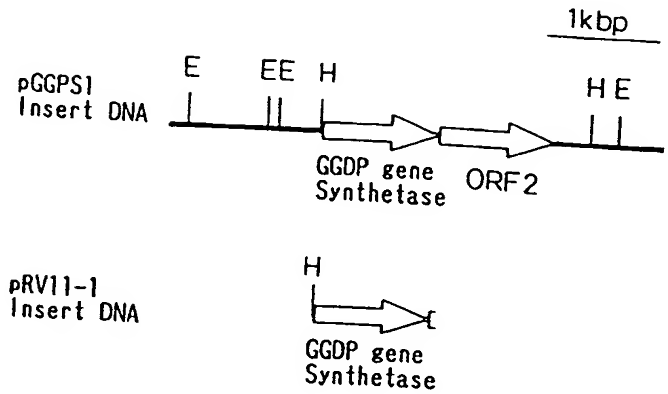
(74) Representative: Tiedtke, Harro, Dipl.-Ing.
Patentanwaltsbüro
Tiedtke-Bühling-Kinne & Partner
Bavariaring 4
D-80336 München (DE)

(54) **Geranylgeranyl diphosphate synthase and DNA coding therefor.**

(57) DNA coding for thermostable geranylgeranyl diphosphate (GGDP) synthase derived from Sulfolobus acidocaldarius is provided. The DNA is useful for production of GGDP synthase, which is, in turn, useful for production of GGDP.

EP 0 674 000 A2

Fig.1


1kbp
pGGPS1
Insert DNA
E
EE
H
H E
GGDP gene
Synthetase
ORF2
pRV11-1
Insert DNA
H
GGDP gene
Synthetase

BACKGROUND OF INVENTION

1. Field of the Invention

The present invention relates to DNA sequence for production of geranylgeranyl diphosphate synthase derived from Sulfolobus acidocaldarius and a transformant with said DNA as well as processes for production of the geranylgeranyl diphosphate synthase and of geranylgeranyl diphosphate using said enzyme.

2. Related Art

Geranylgeranyl diphosphate (GGDP) has four double bonds and includes eight geometrical isomers. GGDP is synthesized in vivo by condensation of isopentenyl diphosphate and farnesyl diphosphate, and is an important intermediate for biosynthesis of isoprenoids and isoprenoid-containing compounds such as calotenoids, diterpenes, vitamines etc. GGDP synthases are found in bacteria, plants, fungi and algae. A large amount of the native isomer of isoprenoids is expressed by introduction by genetic engeneering technique of a gene for GGDP synthase into an appropriate host, and where GGDP synthase is provided in low cost, it can be used for the production of the native isomer of GGDP.

In such a point of view, researches of genes coding for bacterial GGDP synthase and manipulation thereof as well as the production of the synthase have been attempted, and so far only two bacterial genes derived from different sources are known (photosynthetic bacterium Rhodopseudomonos capusulata (J. Bacteriol. 154, p. 580 - 590, 1983; and a phytopathogenic bacterium Erwinia uredovora (J. Bacteriol., 172, p. 6704 - 6712, 1990). These GGDP synthases are unstable, and for example, an enzyme derived from Erwinia uredovora is rapidly inactivated at 55 °C (see, Table 3).

SUMMARY OF THE INVENTION

The unstable enzymes derived from that mesophiles are not sufficient for practical production of GGDP, and especially it is essential to produce a thermostable GGDP synthase. Accordingly, an object of the present invention is to provide primary structure of a gene coding for a thermostable GGDP synthase for developing a process of production of the thermostable GGDP synthase, and to modify microorganisms originally not producing GGDP (such as E. coli) to produce GGDP.

To achieve the above object, we found a GGDP synthase expressed from DNA fragment of Sulfolobus acidocaldarius which is known as an extreme thermophile and acidophilic archaebacterium bacterium. We succeeded to express the GGDP synthase using genetic engineering technique.

The present invention provides a DNA coding for GGDP synthase derived from Sulfolobus acidocaldarius, a recombinant vector having said DNA, as well as recombinant microbial cells to which said gene is introduced by said vector, and the use thereof for the production of GGDP synthase or GGDP per se.

BRIEF EXPLANATION OF DRAWINGS

Fig. 1 represents an insert DNA fragment and a restriction enzyme map thereof in plasmids pGGPS1 and pRVII-1 containing the present DNA sequence. "E" and "H" represent EcoRI and HindIII recognition sites respectively.

Fig. 2 represents a result of analysis of products of a reaction catalyzed by product from plasmid containing DNA of the present invention. Panel A represents a result of thin layer chromatography using LKC-18, and panel C represents a result of that using Kiselgel 60 TLC. Panel B is for a control sample analysed by LKC-18. Circles a, b, c, d and e correspond to geraniol, (all-E) farnesol, (all-E) geranylgeraniol, (2Z, 6E, 10E) geranylgeraniol and (all-E) decaprenol respectively.

Fig. 3 represents thermostability of the present GGDP synthase by determining the remaining enzyme activities. The symbols, solid circle, white circle, cross, solid square, and solid triangle show result of treatment at 60 °C, 70 °C, 80 °C, 90 °C and 100 °C respectively.

Fig. 4 represents the purification steps of the present GGDP as a fusion protein with MBP.

Fig. 5 is a graph showing GGDP synthase activity at each step shown in Fig. 4.

Fig. 6 is a graph showing an enzymatic activity of a Purified MBP-GGDP synthase determined by Grindey-Nichol method.

Fig. 7 is an autoradiogram of TLC showing products formed by reaction of the present fusion protein with various alyllic substrates.

DETAILED DESCRIPTION

According to the present invention, DNA coding for GGDP synthase includes any DNA unit which, on the expression thereof, provide protein having GGDP synthase activity. One particular DNA is that coding for the amino acid sequence shown in SEQ ID NO: 1. The present DNA further include a DNA coding for said amino acid sequence and an additional amino acid sequence (for example, as a fusion protein). A particular example of such a DNA is that having the nucleotide sequence shown in SEQ ID NO: 1.

One embodiment of the present gene is that coding for an amino acid sequence starting with the first Met and ending at 330th Lys in SEQ ID NO: 1. However, there is a case where the first Met is removed by post-translational processing, or the desired enzyme is produced as a fusion protein comprising another peptide. In these cases, the codon coding for the first Met is not present. According to an embodiment of the present invention, the present gene encodes a protein comprising an amino acid sequence starting with the second Ser and ending at the 330th Lys in SEQ ID NO: 1.

An embodiment of the present gene comprises a nucleotide sequence starting with the first nucleotide "A" and ending at 990th nucleotide "A".

It is generally known that same enzymes which are derived from the same species or the same genus involving minor difference in their amino acid sequences by a natural or artificial mutation such as substitution, addition and/or deletion of one or more nucleotides.

According to the invention described in the present specification, it is possible to clone a gene encoding the same enzyme having a difference in amino acid sequence shown in SEQ ID NO: 1. Amino acid sequence drived from such a gene may have a very high homology with the amino acid sequence shown in SEQ ID NO: 1, for example, homology of at least 90%, and further at least 98%. Accordingly, the present invention includes not only a gene coding for GGDP synthase having the amino acid sequence shown in SEQ ID NO: 1, but also gene encoding a protein having a GGDP synthase activity and an amino acid sequence with at least 95% homology, for example 98% homology to that shown in SEQ ID NO: 1.

It is well known for an enzyme that in portions other than portions which is necessary for a function, particular amino acid residues are not essential. Modifications such as substitution, deletion and/or addition of one or a few amino acids can be carried out maintaining the enzyme activity in such non-essential regions. Using general technique such as site-directed mutagenesis with a probe having artificial sequence, we can modify up to 20 amino acids, for example up to 10 amino acids. Substitution, deletion and/or addition of an amino acid sequence can be done using restriction enzymes and/or joining enzymes, ligases.

The present invention includes a gene which encodes a protein having GGDP synthase activity and including an amino acid sequence shown in SEQ ID NO: 1 can be modified by substitution, deletion and/or addition of up to 20 amino acids, for example, up to 10 amino acids.

DNA fragment of the present invention can be prepared according to a procedure per se known as described in detail hereinafter from Sufolobus acidocaldarius available from institutes storing various microorganism, in a form of various length of DNA according to the purpose for using the DNA.

Namely, the present DNA can be prepared by extracting the genomic DNA from Sulfolobus acidocaldarius, cleaving the extracted DNA with, for example, one or more appropriate restriction enzymes to form fragments, inserting the fragments to vectors to prepare a genomic library, and selecting a vector comprising a DNA coding for a desired enzyme by detection of the expression of the desired enzyme. A definite procedure is described in Example 1. a) to d).

Since the present invention discloses a particular nucleotide sequence encoding GGDP synthase, DNA comprising said nucleotide sequence or a nucleotide sequence modified therefrom can be prepared by chemical synthesis. This DNA fragment or a part thereof can be used as primer to synthesize a modified DNA encoding a protein having GGDP synthase activity according to a conventional procedure such as site-directed mutagenesis or PCR method.

The present invention provides recombinant vectors comprising the above-mentioned DNA fragment of the present invention. The recombinant vector can contain a region having functions to express the GGDP synthase gene.

It is known that there are two regulation steps of the gene expression, transcription and translation. Conventional host cell of E. coli has also these regulation systems. As promotor sequences controlling the transcription initiation of mRNA, wild type sequences (for example, lac, trp, bla, lpp, PL, PR, tet, T3, T7 etc.) as well as mutant thereof (for example lacUV5) and artificial fusion sequences of promoter sequences (for example, tac, trc etc.) are known, and can be used in the present invention. The distance between the riposome-binding site (GAGG or similar sequence) and the initiation codon ATG or GTG in some cases are important as a factor which regulates translation of a protein from mRNA. It is well known also that a terminator structure which can stop transcription effects on the efficiency of a recombinant protein

expression (for example, a vector containing rrnBT1T2 is commercially available from Pharmacia).

Vectors which can be used for construction of the present recombinant vectors include commercially available vectors per se, and vectors modified according to purposes. For example, pBR322, pBR327, pKK223-3, pKK233-2, pTrc99 etc. having a replicon derived from pMB1; pUC18, pUC19, pUC118, pUC119, pHSG298, pHSG396, which have been modified to increase the copy number; pACYC177, pACYC184 etc. having a replicon derived from p15A; as well as plasmids derived from pSC101, ColE1, R1, F-factor etc. are mentioned.

Moreover, in addition to plasmids, viral vectors such as λ phage, M13 phage etc., and transpson can be used for introduction of a gene into a host. These vectors are described in Molecular Cloning (J. Sambrook, E.F. Fritsch, T. Maniatis; Cold Spring Harbor Laboratory Press); Cloning Vector (P.H. Pouwels, B.E. Enger•Valk, W.J. Brammer; Elsevier); and various catalogues attached to commercial products.

We can introduce a DNA fragment coding for GGDP synthase and, if necessary, a DNA fragment which can control expression of said enzyme gene into a vector according to known methods using appropriate restriction enzymes and ligases, as described in detail hereinaftar. Plasmids PGGPS1 and pMalGG1 are representative examples of the present plasmids thus constructed.

Microorganisms to be transformed with a recombinant vector thus obtained include Escherichia coli or microorganisms belonging to the genus Bacillus. $CaCl_2$ method, protoplast method etc., as described in, for example, Molecular cloning (J. Sambrook, E.F. Fritsch, T. Maniatis; Cold Spring Harbor Laboratory Press), DNA Cloning Vol. I to III (D.M. Glover; IRL PRESS) etc. can be used for transformation.

A typical transformant of the present invention can be obtained as pGGPS1/DH5α.

We described methods for expression of the desired gene in E. coli above in detail, according to the present invention. A DNA coding for a GGDP synthase can be introduced into other conventional expression vectors according to conventional procedures, such as other prokaryotic cells, lower eukaryotic cells including unicellular host such as yeast, or higher eukaryotic cells such as silk-worm cells. These transformed host cells can be cultured to produce GGDP synthase enzyme.

These transformants or recombinant microbial cells can accumulate GGDP synthase in the cells or in the culture medium while culturing in a medium suitable for said cells such as E. coli. We can prepare GGDP synthase from the cells as follow; lysing the cells with physical disraption or by treatment with a cell-lysing enzyme, removing cell debris to prepare a cell-free extract containing the enzyme, and then isolating and purifying GGDP synthase. We recommend lysozyme as cell-lysing enzyme and sonication as physical disraption. Most of proteins derived from E. coli is denatured by heating at 55°C. The enzyme can be isolated and purified by various chromatographies including gel filtration chromatography, ion exchange chromatography, hydrophobic reversed chromatography, and ultrafiltration and the like alone or in combination. Reducing reagent such as β-mercaptoethanol, dithiothreitol et al., protecting agent against proteases such as PMSF, BSA etc., or metal ions such as magnesium ion can be used to stabilize the desired enzyme during isolation and purification processes, as an enzyme stabilizer.

Activity of GGDP synthase can be determined by, for example, a method described in Example 1. e) It is recommended to isolate and purify GGDP synthase while checking enzyme activity.

The present invention further provides a process for production of GGDP. We can make a host transformed with a DNA encoding GGDP synthase contain DNAs coding for other enzymes in a GGDP biosynthesis pathway. This recombinant can synthesize GGDP by calturing, which can be then prepared and purified.

According to the present invention, the above-mentioned transformant is cultured to produce GGDP synthase, and the isolated enzyme or enzyme-containing product such as partially purified enzyme sample, enzyme-containing cells etc. can be used with substrates, i.e., isopentenyl diphosphate, dimethylallyl diphosphate, geranyl diphosphate or farnesyl diphosphate to synthesize GGDP, which is then recovered.

EXAMPLES

We show primary structure of nucleotide sequence, plasmids and transformants in the following part as well as GGDP synthase and GGDP of the present invention. The present invention is not limited within these Examples.

Example 1.

We carried out procedures with reference to, mainly, the above-cited Molecular Cloning, and DNA Cloning as well as catalogues from Takara Shuzo. Most of enzymes were purchased from Takara Shuzo. Reversed phase thin layer chromatography (TLC) plates LKC-18 were purchased from Whatman, Kieselgel

60 thin-layer chromatography (TLC) plates LKC-18 were purchased from Merck. Sulfolobus acidocaldarius ATCC 33909 used in the present invention is registered in and available without any limitation from America Type culture Collection (ATCC).

a) Preparation of chromosomal DNA from Sulfolobus acidocaldarius

Cells of ATCC 33909 strain were cultured in the 1723 medium described in ATCC catalogue. Genomic DNA was prepared from the cultured cells, according to Current Protocols in Molecular Biology published by Wiley Interscience.

b) Preparation of a genomic DNA library of Sulfolobus acidocaldarius

The chromosomal DNA was partially digested with a restriction enzyme Sau3AI, and subjected to 0.5% agarose gel electrophoresis. A block of the agarose gel containing DNA fragments of 3 kbp to 6 kbp was fractionated, and DNA was extracted therefrom. 2.7 µg of this size-fractionated DNA and 1.4 µg of pUC119 plasmid DNA cleaved with BamHI and dephosphorylated were ligated using DNA ligase. The ligation cocktail was used to transform E. coli DH5α. The transformants were then stored at -70°C. The library thus constructed was screened.

c) Preparation of competent cells carrying plasmid pACYC-IB

A 2.8 kbp of SnaBI-HpaI DNA fragment containing crtI (phytoene synthase gene) and crtB (phytoene desaturase gene) of Erwinia uredovora was prepared from pCAR25 (N. Misawa et at., J. Bacterial. 172. 6704 - 6712 (1990)). Erwinia uredovora is avaialble from ATCC19321. It is easily grown wtih LB medium. That 2.8 kbp DNA fragment can be cloned, for instance, using popular tehcniqus of Southern analysis or PCR with probe DNAs synthesized from crtI or crtB nucleotide sequences retrievable from GenBank accession No. D90087.

This DNA fragment was ligaled with an EcoRI linkers and cleaved with a restriction enzyme EcoRI, and joined using DNA ligase to a plasmid pACYC184 which had been digested with EcoRI and dephosphorylated. The ligation cocktail was then used to transform E. coli pACYC-IB/pH5α. Competent calls were prepared by $CaCl_2$ method.

d) Selection of GGDP synthase gene

Plasmid DNAs were purified from E. coli containing Sulfolobus acidocaldarius genomic DNA library constructed in the above-mentioned procedure b) according to an alkaline method, and 10 nanograms of the DNA was used to transform the E. coli carrying the crtI and crtB genes prepared by the above-mentioned method c). The transformants were cultured on an LB agar plates containing 50 µg/ml tetracycline and 50 µg/ml ampicillin.

10 positive red colonies were obtained by visually selecting among about 4,000 transformants. Among these 10 colonies a plasmid named pGGPS1 was isolated from one colony. A 2.3 kbp of HindIII fragment in pGGPS1 insert DNA was subcloned into pUC118. This subclone DNA was then introduced into E. coli cells carrying the crtI and crtB gene. This clone prduced red colonies. Since the presence of GGDP synthase gene in the HindIII fragment was demonstrated, nucleotide sequence of this 2.3 kbp was determined by dideoxy chain termination method.

As a result, the 2.3 kbp fragment included two open reading frames (ORF-1 and ORF-2) as shown in Fig. 1.

Accordingly, a plasmid pRV11-1 lacking the downstream ORF-2 was constructed, and enzyme activity of an expression product was determined according to the method e).

e) Assay of GGDP synthase activity

The plasmid pRV11-1 obtained in the above-mentioned method d) was used to transform E. coli DH5α, which was then cultured in 100 ml of LB medium containing 50 µg/ml ampicillin at 37°C overnight. The cells were harvested and disrupted by sonication in 8 ml of Sonic buffer (10 mM 2-mercaptoethanol, 1 mM EDTA, 50 mM Tris-HCl (pH7)), and the homogenate was heated at 55°C for 60 minutes and centrifuged at 10,000 x g for 10 minutes. The supernatant was used to assay for GGDP synthase activity.

The assay mixture contained, in a final volume of 1 ml, 0.48 μ mol of [1-^{14}C] isopentenyl diphosphate (1.92 GBg/mmol), 25 μ mol of (all-E) farnesyl diphosphate, 5 μ mol of $MgCl_2$, 25 μ mol of Tris-HCl (pH6.8) and 0.3 mg the above crude enzyme. This mixture was incubated at 55°C for 30 minutes, and chilled in an ice bath to stop the reaction. The reaction mixture was extracted with 3 ml of 1-butanol saturated with water, and the radioactivity in the 1-butanol layer was counted to determine GGDP synthase activity. A result is shown in Table 1. It was shown that all of the clones thus obtained have a gene which was expected to encode a thermostable GGDP synthase. In addition, an assay of an extract from the clone containing pRV11-1 showed that the ORF-1 is GGDP synthase gene.

Table 1

Result of assay for GGDP synthase activity derived from plasmid of the present invention (Radioactivity of 1-butanol extract is shown in dpm unit.)

Cell-free extract from:	Enzyme activity (dpm)
E. coli DH5α (no heat treatment)	3,310
E. coli DH5α (heat treatment)	0
E. coli DH5α/pGGPS1 (heat treatment)	11,900
E. coli DH5α/pRV11-1 (heat treatment)	8,770

f) Analysis of GGDP synthase product

Identification of a product produced by the heat-denatured cell-free extract was carried out.

The products obtained from the incubation of [1-^{14}C] isopentenyl diphosphate and furnesyl diphosphate with a cell-free extract from selected positive transformant were hydrolyzed with an acid phosphatase according to a method of Fujii et al. (Fujii et al., (1982) Biochim. Biophys. Acta 712, p. 716 - 718). The hydrolyzed alcohols were extracted with pentane. The pentane-soluble products were analyzed by reversed phase LKC-18 thin layer chromatography using a mixed solvent of acetone/water (9:1) and normal phase Kieselgel 60 thin layer chromatography using a mixed solvent of benzene/ethyl acetate (9:1). A result is shown in Fig. 2. It is clearly shown that the radio-active alcohol derived from the recombinant product is (all-E) geranylgeraniol which is a derivative from (all-E)-GGDP.

Accordingly, successful cloning of GGDP synthase gene from Sulfolobus acidocaldarius was confirmed.

g) Partial purification of GGDP synthase derived from cloned gene

A heat-treated cell-free extract from a cell lysate of recombinant E. coli carrying a GGDP synthase gene was precipitated with 30 - 60% saturation of $(NH_4)_2SO_4$. The precipitated protein fraction was dialyzed and chromatographed on a DEAE Toyopearl 650M column (1.0 × 16 cm) equilibrated with buffer A (1 mM EDTA, 10 mM Tris-HCl (ph7.7)); Elution was performed with a linear gradient from 0 to 0.85M NaCl in buffer A. Fractions containing GGDP synthase were collected and dialyzed against buffer A. The dialysate was applied to a Mono Q column (5 × 50 mm) equilibrated with buffer A; Elution was performed with a linear gradient of 0 - 0.85M of NaCl in buffer A. A fraction containing GGDP synthase was analysed by 10% SDS polyacrylamide gel electrophoresis and the gel was stained with Coamassie Briliant Blue.

After these operations, a specific activity reached 8.7 nmol/min./mg protein.

h) Substrate specificity of GGDP synthase derived from cloned gene

Substrate specificity of GGDP synthase derived from the cloned gene was tested using allyl diphosphate substrates shown in Table 2. As a result, it was found that dimethylallyl diphosphate, geranyl diphosphate and (all-E) farnesyl diphosphate can be substrates.

Table 2

Substrate specificity of GGDP synthase derived from cloned gene of the present invention	
Substrate	Enzyme activity (dpm)
Dimethylallyl diphosphate	24,900
Geranyl diphosphate	20,900
(all-E) Farnesyl diphosphate	15,300
(2Z, 6E) Farnesyl diphosphate	260
(all-E) Geranylgeranyl diphosphate	0
(2Z, 6E, 10E) Geranylgeranyl diphosphate	50

i) Thermostability of GGDP synthase derived from cloned gene

The GGDP synthase derived from the gene cloned from Sulfolobus acidocaldarius was heat-treated and remaining activity was determined. After heating at 60°C for 100 minutes, at least 95% of activity was maintained. A result is shown in Fig. 3.

For comparison, thermostability of GGDP synthase derived from Erwinia uredovora (product of crtE) was as follow.

Table 3

Thermostability of GGDP synthase derived from Erwinia uredovora (crtE product)	
Combination of treatment	Remaining activity
55°C 5 min.	55%
55°C 10 min.	47%
55°C 30 min.	3%

Example 2. Production of GGDP synthase

A polymerase chain reaction (PCR) was carried out using the above-mentioned plasmid pGGPS1 as a template, the following primers:

GGPP-I BamHI (26 mer. 5'-CGC CGA TCC ATG AGT TAC TTT GAC AA-3' (SEQ ID NO: 2), and

GGPP-T EcoRI (25 mer. 5'-GG GAA TTC TTA TTT TCT CCT TCT TA-3') (SEQ ID NO: 3),

and the reaction composition shown in Table 4, to amplify a DNA fragment corresponding to the coding region of GGDP synthase gene of the present invention.

Table 4

Reaction composition of PCR	
Template DNA	1.0 µl
10 pmol/µl primer 1 (GGPS-I BamHI)	1.0 µl
10 pmol/µl primer 2 (GGPS-T EcoRI)	1.0 µl
dNTP mix (Takara Shuzo)	4.0 µl
AmpliTaq (5 U/µl)	1.0 µl
10x AmpliTaq buffer (Takara Shuzo)	10.0 µl
1 u/µl perfect match polymerase enhancer (Stratagene)	1.0 µl
Distilled water	81.0 µl
Total	100.0 µl

The PCR condition was 30 cycles of 90°C for 30 seconds, 50°C for 30 seconds and 72°C for one minute. After finishing the reaction, the amplified DNA was precipitated with ethanol at -80°C, cleaved with a restriction enzyme EcoRI, blunt-ended, and further cleaved with a restriction enzyme BamHI to obtain a DNA fragment of about 1 kbp coding for GGDP synthase.

A commercially available plasmid pMAL-c2 (NEB, USA) was used as a cloning expression vector. In this plasmid, a DNA fragment (the name of gene: malE) coding for maltose-binding protein (sometimes designated MBP hereinafter) is inserted downstream of tac promoter, and a cloning sites for desired gene (DNA) is present downstream of the malE. Accordingly, where a gene encoding a desired polypeptide is inserted into the cloning site and the gene is expressed, then a fusion protein of the MBP and the desired polypeptide is formed, and the fusion protein can be purified in a single step by an amylose resin affinity chromatography.

The plasmid pMAL-c2 was cleaved with a restriction enzyme HindIII, blunt-ended and cleaved with a restriction enzyme BamHI. The resulting DNA fragment was ligated with the PCR-amplified DNA fragment coding for the GGDP synthase so as to obtain a recombinant plasmid pMalcGG1. The ligation and blunt-ending were carried out using a ligation kit and blunting kit of Takara Shuzo.

This recombinant plasmid was used to transform E. coli TOPP cell NO. 2 (Stratagene) (competent cell). The transformed cells were cultured on a YT plate medium resulting in formation of 6 colonies. These colonies were cultured in 2xYT liquid medium at 37°C, and plasmid DNA was prepared. We tested to confirm that the recombinant plasmid was correctly constructed by cleavage with EcoRV, and checking size of DNA bands by electrophoresis. Note that the correct recombinant plasmid, when cleaved with EcoRV, provides two DNA fragments of 5.1 kbp and 2.5 kbp.

The plasmid pMalcGG1 was used to transform E. coli pACYC-IB/DH5α (Ohnuma at al., J. Biol Chem. 1994; 269 (20): 4792 - 4791) to obtain a transformant pMalcGG1, pACYC-IB/DH5α. Note, the E. coli pACYC-IB/DH5α already has plasmid pACYC-IB, and the pACYC-IB contains enzymes which joins two geranyl-geranyl diphosphate (GGDP) (the number of carbon atoms: 20) to form phytoene, and further contains gene for desaturation, and expresses these genes.

The transformant was cultured over night in 100 ml of LB medium, and then was inoculated to 1L of LB medium, and cultured at 37°C under the stirring conditions at 300 rpm.

When cell concentration reached to kelett = 30 to 40, 10 ml of 100 mM IPTG was added to the culture to induce the transcription, and culturing was further carried out for 4 hours. As a control, a sample was taken immediately before the addition of IPTG. The culture was centrifuged to collect the cells, which were then disrupted by sonication. White protein samples after induction of expression with IPTG and before induction of expression with IPTG were analyzed by SDS-polyacrylamdie gel electrophoresis (SDS-page) as shown in Fig. 4 wherein IPTG (+) and IPTG (-) were referred to respectively. As a result, it was confirmed that fusion protein of about 70 kb was abundantly present in the IPTG (+) lane. SDS-page was performed according to Wiley et al., Current Protocols in Molecular Biology and using a miniproteo-II cell apparatus of Bio Rad, and gel was stained with Coomassie Brilliant Blue and dried using a gel drying kit of Promega.

The above-mentioned cell disruptant was centrifuged to fractionate into a supernatant and a precipitation. An analysis by SDS-PAGE for these fractions are shown in Fig. 4 as "Sonication sup." and "Sonication ppt." respectively. It is could be found that the fusion protein of about 70 kb was transferred to the supernatant. The supernatant was heated at 60°C for one hour, and denatured protein was removed by centrifugation so as to obtain a supernatant. A result of analysis therefor is shown in Fig. 4 as "Sonication sup. (h+)". It is confirmed that impurity was removed and the fusion protein was enriched.

Next, the fusion protein was purified by affinity chromatography. For the purification, the above-mentioned supernatant was filtered through a 0.45 µm or 0.20 µm membrane filter, and the filtrate was passed through a column (2.5 x 10 cm) filled with 15 ml of amylose resin, and elution was carried out according to a protocol of NEB attached to the plasmid pMAL-c2. The eluant was desalted with PD-10 column (Pharmacia), to obtain about 3.4 mg of fusion protein. The fusion protein wad heated at 60°C for one hour. A result of SDS-PAGE for the fusion proteins prior heating and after heating is shown us "Fusion" and "Fusion (h+)" respectively. It is confirmed that the fusion protein was not denatured by heating.

Next, the fusion protein thus obtained was cleaved with Factor Xa to liberate GGDP synthase. A result of SDS-PASE for the cleavage product is shown in Fig. 4 as "Fusion (digested)", and for the cleavage product heated at 60°C for one hour as "Fusion (digested) (h+)". As a result, it was confirmed that GGDP synthase was partially liberated.

In addition, for each fraction described above, GGDP synthase activity was measured by method described in Exmple 1. e). This assay was carried out as follow. A solution containing 2 µg of protein, 5 mM $MgCl_2$, 10 mM KH_2PO_4/KOH (pH5.8), 25 µM substrate (geranye diphosphate, (all-E)-farnesyl diphosphate, or (2Z, 6E)-furnesyl diphosphate), and 463 nM [^{14}C]-isopentenyl diphosphate (4 Ci/mole)/ml was reacted at

55 °C for one hour, and the reaction mixture was extracted with 3 ml water-saturated butanol. One ml of the extract was used for assay of radioactivity, and the remaining portion was hydrolyzed with potato acid phosphatase, extracted, and analyzed by TLC. A result is shown in Fig. 5.

Plasmid pMalcGG1 was used to transform E. coli JM105 so as to obtain a transformant pMalcGG1/JM105. The transformant was cultured in TYGPN medium (20g trypton, 10g yeast extract, 10 ml 80% glycerol, 5g Na_2HPO_4, 10g KNO_3/L) for 4 hours (klett = 32), induced the expression by IPTG, and further cultured for 26 hours. MBP-GGDP synthase fusion protein was purified as described above so as to obtain 21.8 mg of purified fusion protein per 1L. Note, the broth in which the MBP-GGDP synthase fusion protein wag produced was red to deep red color, and this recombinant cell was redish brown in comparison with broth and cell in which MBP alone was expressed (pMAL-c2/JM109).

GGDP synthase activity of the MBP-GGDP synthase fusion protein could be also determined by measuring inorganic phosphate according to Grindey-Nichol method (Grindy & Nichol, Anal. Biochem. 1970, 33, 114 - 119). A reaction mixture containing 50 mM Tris•HCl, 5 mM $MgCl_2$, 50 mM NH_4Cl, 10 mM 2-mercaptoethanol, 50 nmole farnesyl diphosphate, 50 nmole isopentenyl diphosphate, and for example 200 µg enzyme sample to be tested was reacted at 55 °C for 3 hours, and cooled to 0 °C to stop the reaction. Protein in the reaction mixture was denaturated with trichloroacetic acid and removed, and the supernatant was neutralized with NaOH, and amounts of orthophosphate and pyrophosphate in the reaction mixture were measured to determine enzyme activity. A result is shown in Fig. 6.

Note that when GGDP synthase coding region in the expression plasmid pMalcGG1 was sequenced, the 720th nucleotide "A" was substituted by "G", and the 740th nucleotide "A" was changed to "G", resulting in change of the 247th amino acid Lys to Arg. It is assumed that the changes occurred during the PCR.

In addition, when the same experiment as described above was repeated using pGEX-2T (Pharmacia) in place of pMAL-c2, in the resulting plasmid pGluTGG1, the 823th nucleotide "A" was changed to "G", resulting in change of the 275th amino acid Met to Val. As a result, a glutathione S-transferase GGDP synthase fusion protein having a point mutation at the 275th position was obtained. This fusion protein derived from pGluTGG1 exhibited the enzyme activity as the same enzyme derived from pMalcGG1.

On the other hand, where pGEX-3X (Pharmacia) was used in place of pMAL-c2 as an expression cloning plasmid, in the resulting pGluXGG1, no mutation occurred, and glutathione S-transferase GGDP synthase fusion protein having the amino acid sequence shown in SEQ ID NO: 1 was obtained. This pGluXGG1-derived fusion protein also exhibited the same enzyme activity as the pMalcGG1-derived fused protein.

Example 3. Production of geranylgeranyl compound

Fusion proteins which were expression products of the above-mentioned expression products pMalcGG1, pGluTGG1 and pGluXGG1 as well as cleavage products (digestion products) thereof were reacted with substrates (primers), i.e., geranyldiphosphate (GPP), (all-E) farnesyl-diphosphate ([All-E]-FPP), and (2Z, 6E) furnesyl diphosphate ([2Z, 6E]-FPP). Resulting products are shown in Fig. 7.

In this figure, "C" represents a protein derived from pGluTGG1, "D" represents a protein derived from pGluXGG1, and "E" represents a protein derived from pMalcGG1. "(f)" represents a fused protein purified with an affinity column, and "(d)" represents a reaction product affinity-purified and treated with thrombin for pGluTGG1, and a product affinity-purified and treated with Factor-Xa for pGluXGG1 and pMalcGG1.

According to the present invention, DNA sequence coding for GGDP synthase derived from Sulfolobus acidocaldarius is provided. Recombinant cells, such as E. coli cells, transformed with an expression plasmid containing said DNA fragment produce stable, especially thermostable enzyme having geranylgeranyl diphosphate activity and geranylgeranyl diphosphate.

SEQUENCE LISTINGS

SEQ ID NO: 1

Sequence Length: 993

Sequence Type: Nucleic acid

Strandness: Double

Topology: Linear

Molecular type: Genomic DNA

Source

Organism: Sulfolobus acidocaldarius

Sequence:

```
ATG AGT TAC TTT GAC AAC TAT TTT AAT GAG ATT GTT AAT TCT GTA      45
Met Ser Tyr Phe Asp Asn Tyr Phe Asn Glu Ile Val Asn Ser Val
                  5                  10                  15
AAC GAC ATT ATT AAG AGC TAT ATA TCT GGA GAT GTT CCT AAA CTA      90
Asn Asp Ile Ile Lys Ser Tyr Ile Ser Gly Asp Val Pro Lys Leu
                 20                  25                  30
TAT GAA GCC TCA TAT CAT TTG TTT ACA TCT GGA GGT AAG AGG TTA     135
Tyr Glu Ala Ser Tyr His Leu Phe Thr Ser Gly Gly Lys Arg Leu
                 35                  40                  45
AGA CCA TTA ATC TTA ACT ATA TCA TCA GAT TTA TTC GGA GGA CAG     180
Arg Pro Leu Ile Leu Thr Ile Ser Ser Asp Leu Phe Gly Gly Gln
                 50                  55                  60
AGA GAA AGA GCT TAT TAT GCA GGT GCA GCT ATT GAA GTT CTT CAT     225
Arg Glu Arg Ala Tyr Tyr Ala Gly Ala Ala Ile Glu Val Leu His
                 65                  70                  75
ACT TTT ACG CTT GTG CAT GAT GAT ATT ATG GAT CAA GAT AAT ATC     270
Thr Phe Thr Leu Val His Asp Asp Ile Met Asp Gln Asp Asn Ile
                 80                  85                  90
AGA AGA GGG TTA CCC ACA GTC CAC GTG AAA TAC GGC TTA CCC TTA     315
Arg Arg Gly Leu Pro Thr Val His Val Lys Tyr Gly Leu Pro Leu
                 95                 100                 105
GCA ATA TTA GCT GGG GAT TTA CTA CAT GCA AAG GCT TTT CAG CTC     360
Ala Ile Leu Ala Gly Asp Leu Leu His Ala Lys Ala Phe Gln Leu
                110                 115                 120
```

TTA ACC CAG GCT CTT AGA GGT TTG CCA AGT GAA ACC ATA ATT AAG 405
Leu Thr Gln Ala Leu Arg Gly Leu Pro Ser Glu Thr Ile Ile Lys
125 130 135

GCT TTC GAT ATT TTC ACT CGT TCA ATA ATA ATT ATA TCC GAA GGA 450
Ala Phe Asp Ile Phe Thr Arg Ser Ile Ile Ile Ile Ser Glu Gly
140 145 150

CAG GCA GTA GAT ATG GAA TTT GAG GAC AGA ATT GAT ATA AAG GAG 495
Gln Ala Val Asp Met Glu Phe Glu Asp Arg Ile Asp Ile Lys Glu
155 160 165

CAG GAA TAC CTT GAC ATG ATC TCA CGT AAG ACA GCT GCA TTA TTC 540
Gln Glu Tyr Leu Asp Met Ile Ser Arg Lys Thr Ala Ala Leu Phe
170 175 180

TCG GCA TCC TCA AGT ATA GGC GCA CTT ATT GCT GGT GCT AAT GAT 585
Ser Ala Ser Ser Ser Ile Gly Ala Leu Ile Ala Gly Ala Asn Asp
185 190 195

AAT GAT GTA AGA CTG ATG TCT GAT TTC GGT ACG AAT CTA GGT ATT 630
Asn Asp Val Arg Leu Met Ser Asp Phe Gly Thr Asn Leu Gly Ile
200 205 210

GCA TTT CAG ATT GTT GAC GAT ATC TTA GGT CTA ACA GCA GAC GAA 675
Ala Phe Gln Ile Val Asp Asp Ile Leu Gly Leu Thr Ala Asp Glu
215 220 225

AAG GAA CTT GGA AAG CCT GTT TTT AGT GAT ATT AGG GAG GGT AAA 720
Lys Glu Leu Gly Lys Pro Val Phe Ser Asp Ile Arg Glu Gly Lys
230 235 240

AAG ACT ATA CTT GTA ATA AAA ACA CTG GAG CTT TGT AAA GAG GAC 765
Lys Thr Ile Leu Val Ile Lys Thr Leu Glu Leu Cys Lys Glu Asp
245 250 255

GAG AAG AAG ATT GTC CTA AAG GCG TTA GGT AAT AAG TCA GCC TCA 810
Glu Lys Lys Ile Val Leu Lys Ala Leu Gly Asn Lys Ser Ala Ser
260 265 270

AAA GAA GAA TTA ATG AGC TCA GCA GAT ATA ATT AAG AAA TAC TCT 855
Lys Glu Glu Leu Met Ser Ser Ala Asp Ile Ile Lys Lys Tyr Ser
275 280 285

TTA GAT TAT GCA TAC AAT TTA GCA GAG AAA TAT TAT AAA AAT GCT 900
Leu Asp Tyr Ala Tyr Asn Leu Ala Glu Lys Tyr Tyr Lys Asn Ala
290 295 300

ATA GAC TCT TTA AAT CAA GTC TCC TCT AAG AGT GAT ATA CCT GGA 945
Ile Asp Ser Leu Asn Gln Val Ser Ser Lys Ser Asp Ile Pro Gly
305 310 315

AAG GCT TTA AAA TAT CTA GCT GAA TTT ACG ATA AGA AGG AGA AAA 990
Lys Ala Leu Lys Tyr Leu Ala Glu Phe Thr Ile Arg Arg Arg Lys
320 325 330

TAA 993
TER

SEQ ID NO: 2
Sequence Length: 26
Sequence Type: Nucleic acid
Strandness: Single
Topology: Linear
Molecular type:
Sequence:
CGCGGATCCA TGAGTTACTT TGACAA 26

SEQ ID NO: 3
Sequence Length: 25
Sequence Type: Nucleic acid
Strandness: Single
Topology: Linear
Molecular type:
Sequence:
GGGAATTCTT ATTTTCTCCT TCTTA 25

DNA coding for thermostable geranylgeranyl diphosphate (GGDP) synthase derived from Sulfolobus acidocaldarius is provided. The DNA is useful for production of GGDP synthase, which is, in turn, useful for production of GGDP.

Claims

1. DNA coding for geranylgranyl diphosphate synthase (GGDP synthase) of Sulfolobus acidocaldarius origin.

2. DNA according to claim 1, wherein the enzyme consists essentially of the amino acid sequence shown in SEQ ID NO: 1.

3. DNA according to claim 1, wherein the DNA consists essentially of the nucleotide sequence shown in SEQ ID NO: 1.

4. A recombinant vector comprising DNA according to claim 1 and DNA region having a function to regulate the expression of said DNA.

5. Recombinant microbial cell which is transformed with a recombinant vector according to claim 4.

6. Recombinant microbial cell according to claim 5, wherein the host belongs to the genus Escherichia.

7. A process for production of GGDP synthase comprising the steps of culturing the recombinant microbial cell according to claim 5 in a medium, and recovering a GGDP synthase-active substance.

8. A process for production of geranylgranyl diphosphate (GGDP) comprising the steps of culturing recombinant microbial cell according to claim 5 in a medium, and recovering GGDP.

9. A process for production of GGDP comprising reacting a culture of the recombinant microbial cell according to claim 5 with a substrate selected from the group consisting of isopentenyl diphosphate, dimethylallyl diphosphate, geranyl diphosphate and farnesyl diphosphate.

10. A process for production of GGDP, comprising reacting an enzyme-active substance obtained by a process according to claim 7 with a substrate selected from the group comprising of isopentenyl diphosphate, dimethylallyl diphosphate, geranyl diphosphate and farnesyl diphosphate.

Fig.1

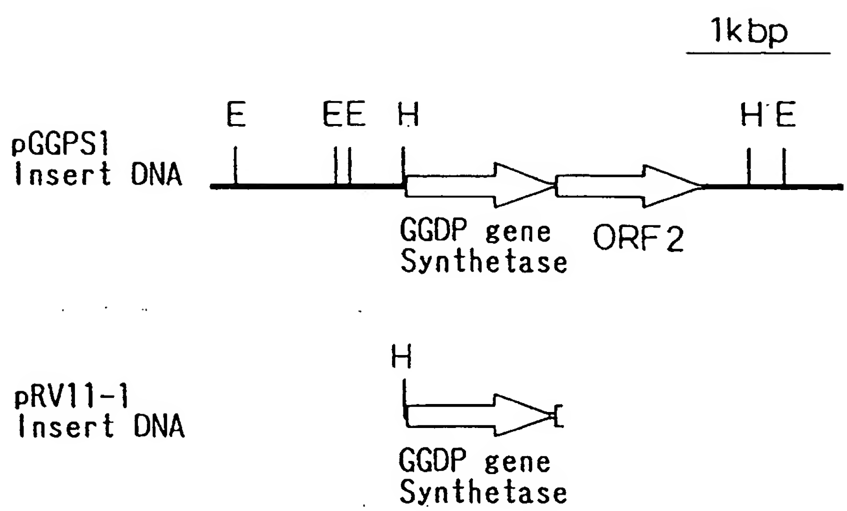

1kbp
E
EE
H
H E
pGGPS1
Insert DNA
GGDP gene
Synthetase
ORF2
H
pRV11-1
Insert DNA
GGDP gene
Synthetase

Fig.2

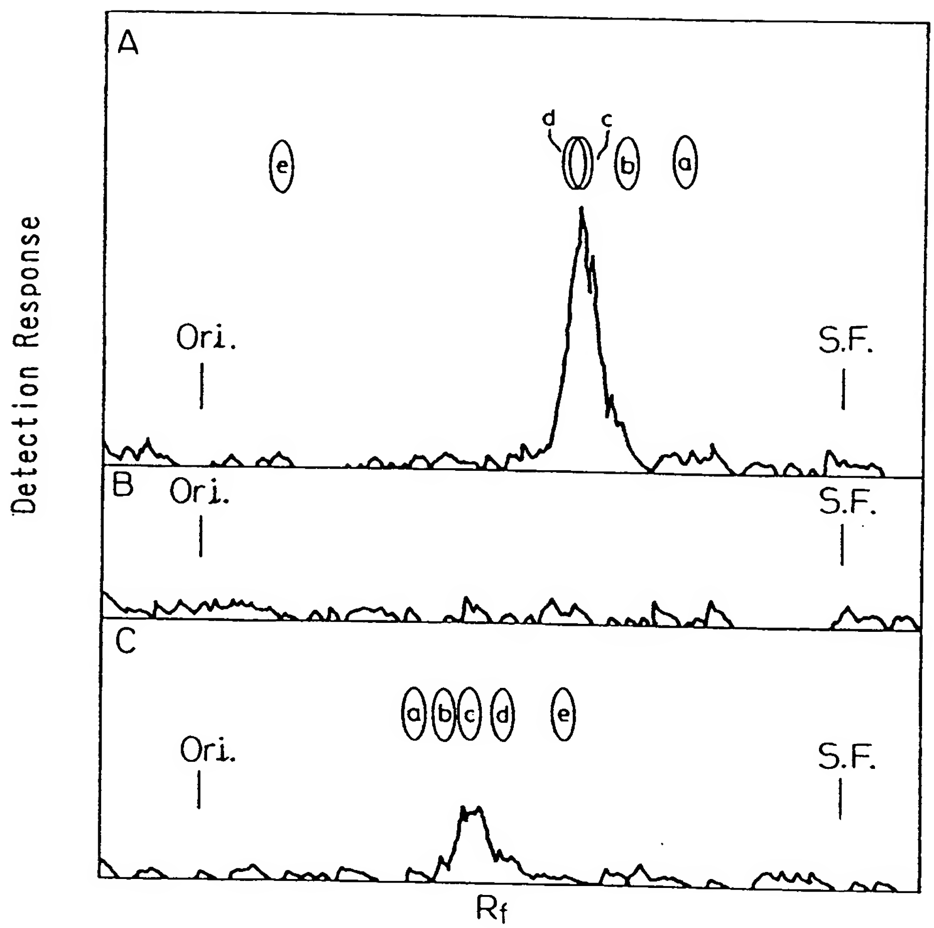


Fig.3

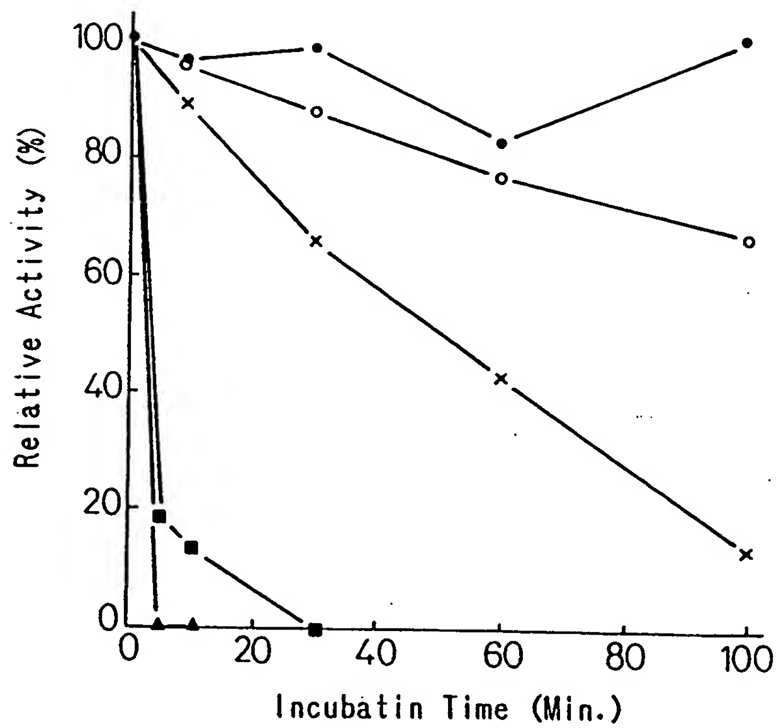

100
80
60
40
20
0
Relative Activity (%)
0
20
40
60
80
100
Incubatin Time (Min.)

Fig.4

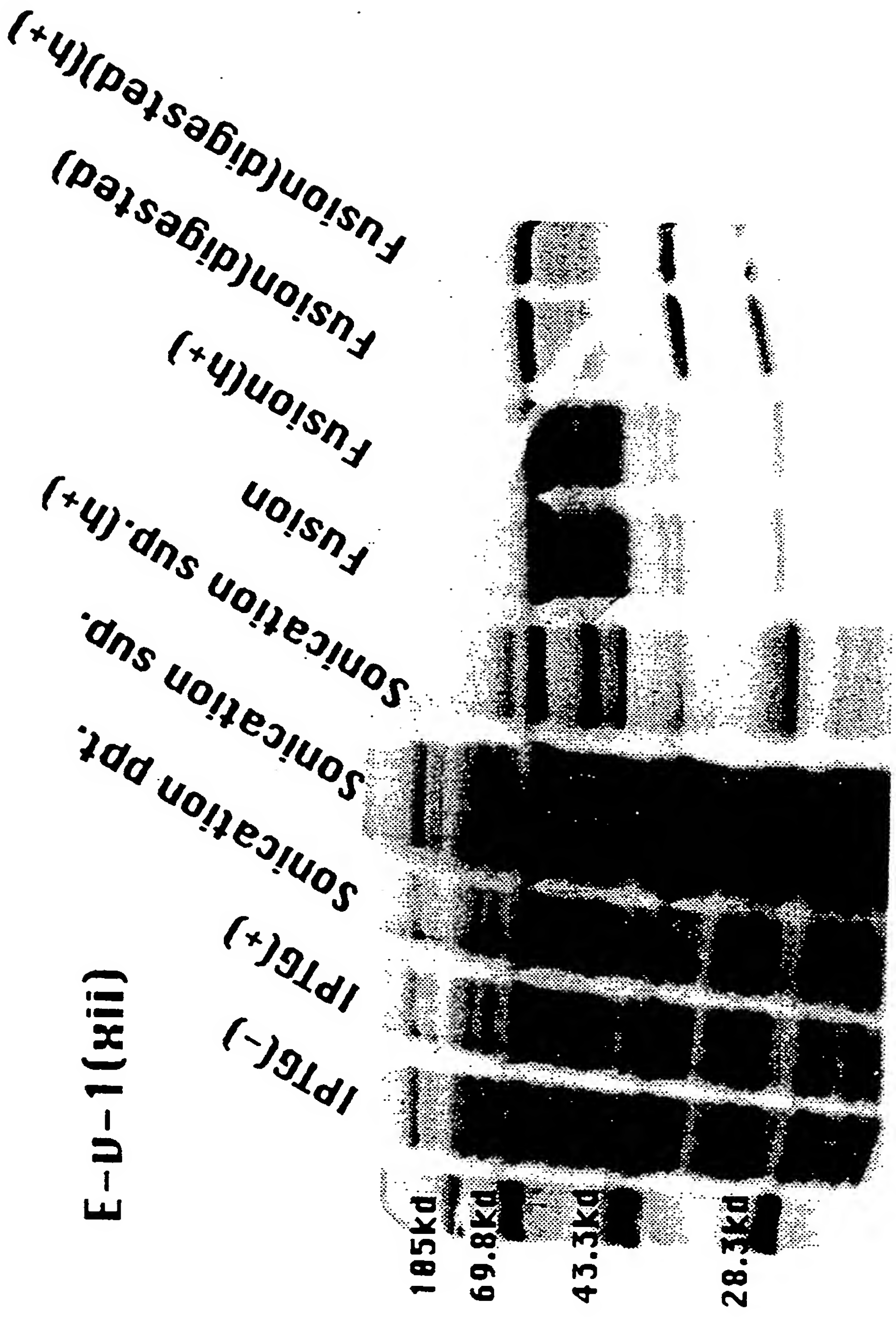

E-U-1(xiii)
IPTG(-)
IPTG(+)
Sonication ppt.
Sonication sup.
Sonication sup.(h+)
Fusion
Fusion(h+)
Fusion(digested)
Fusion(digested)(h+)
185kd
69.8kd
43.3kd
28.3kd

Fig.5

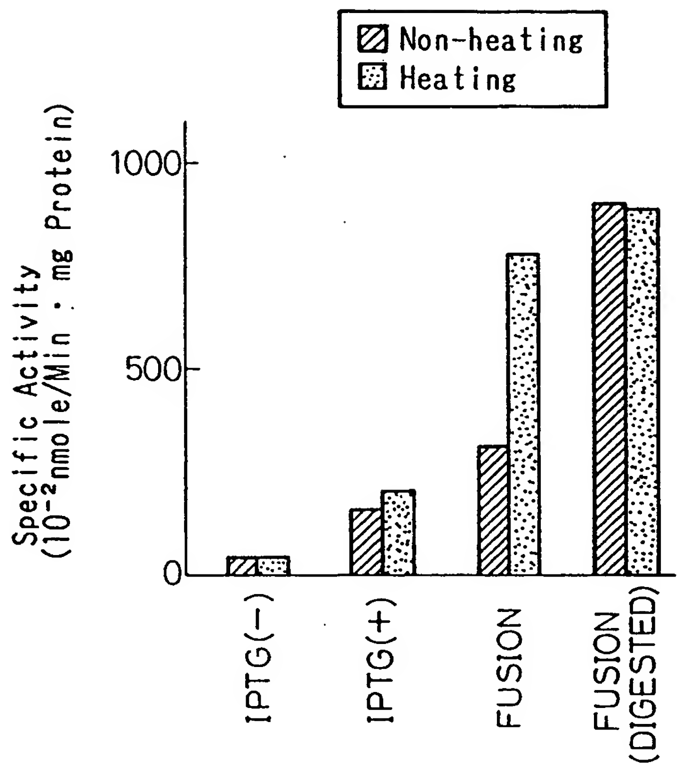

Non-heating
Heating
Specific Activity
(10^{-2}nmole/Min · mg Protein)
1000
500
0
IPTG(−)
IPTG(+)
FUSION
FUSION
(DIGESTED)

Fig.6

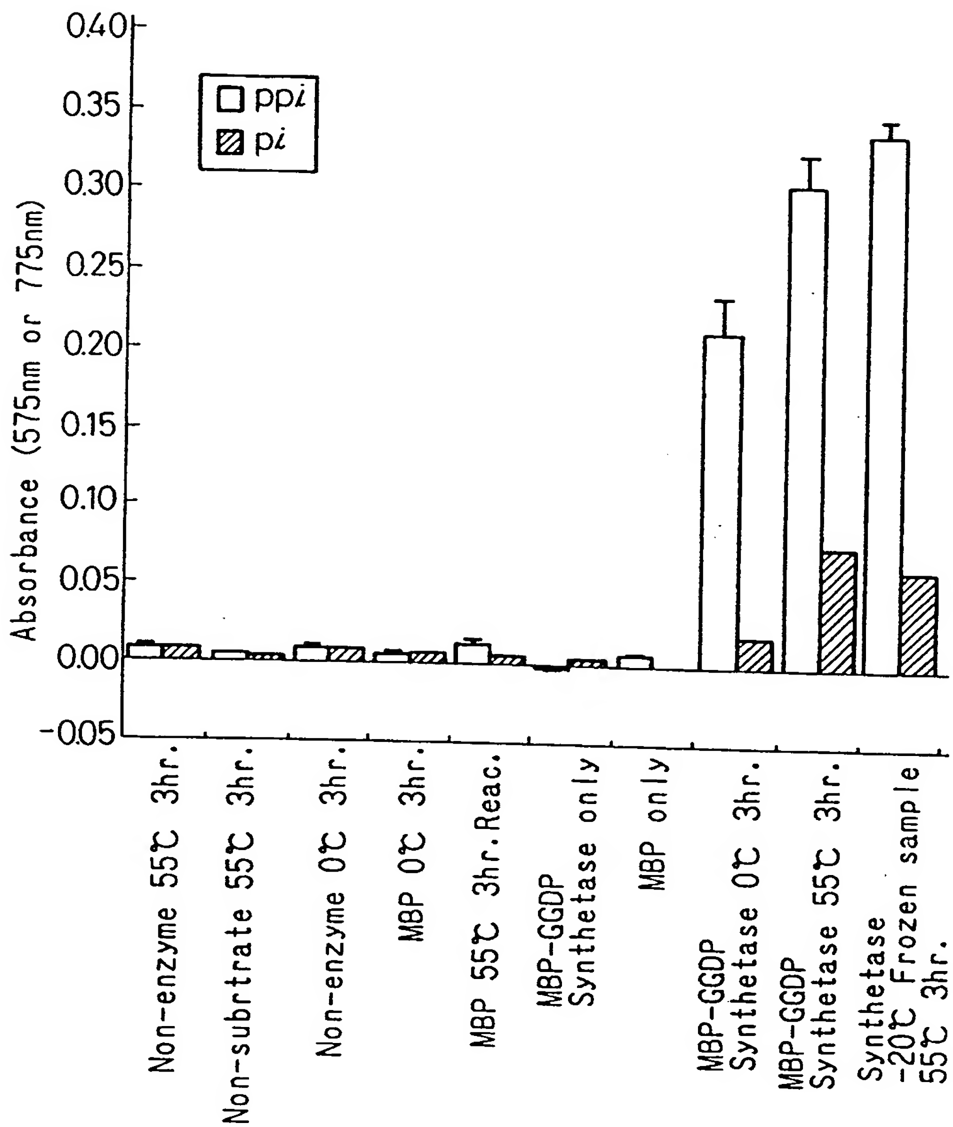

0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
-0.05
Absorbance (575nm or 775nm)
ppi
pi
Non-enzyme 55℃ 3hr.
Non-subrtrate 55℃ 3hr.
Non-enzyme 0℃ 3hr.
MBP 0℃ 3hr.
MBP 55℃ 3hr. Reac.
MBP-GGDP Synthetase only
MBP only
MBP-GGDP Synthetase 0℃ 3hr.
MBP-GGDP Synthetase 55℃ 3hr.
Synthetase -20℃ Frozen sample 55℃ 3hr.

Fig.7


F_OH →
GG_OH →
C(r) D(r) E(r) C(d) D(d) E(d)
primer:GPP
C(r) D(r) E(r) C(d) D(d) E(d)
primer:(all-E)-FPP
C(r) D(r) E(r) C(d) D(d) E(d)
primer:(2Z,6E)-FPP

Fig.1

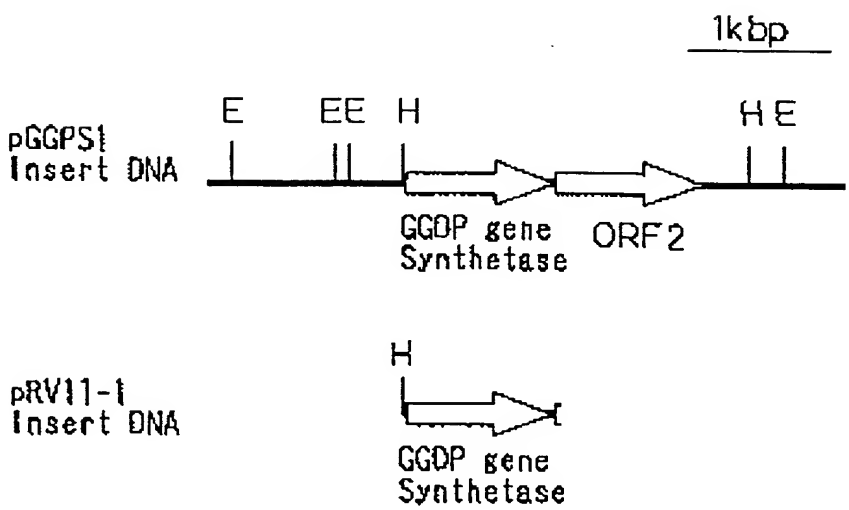

1kbp
E
EE
H
H E
pGGPS1
Insert DNA
GGDP gene
Synthetase
ORF2
H
pRV11-1
Insert DNA
GGDP gene
Synthetase

Fig.2

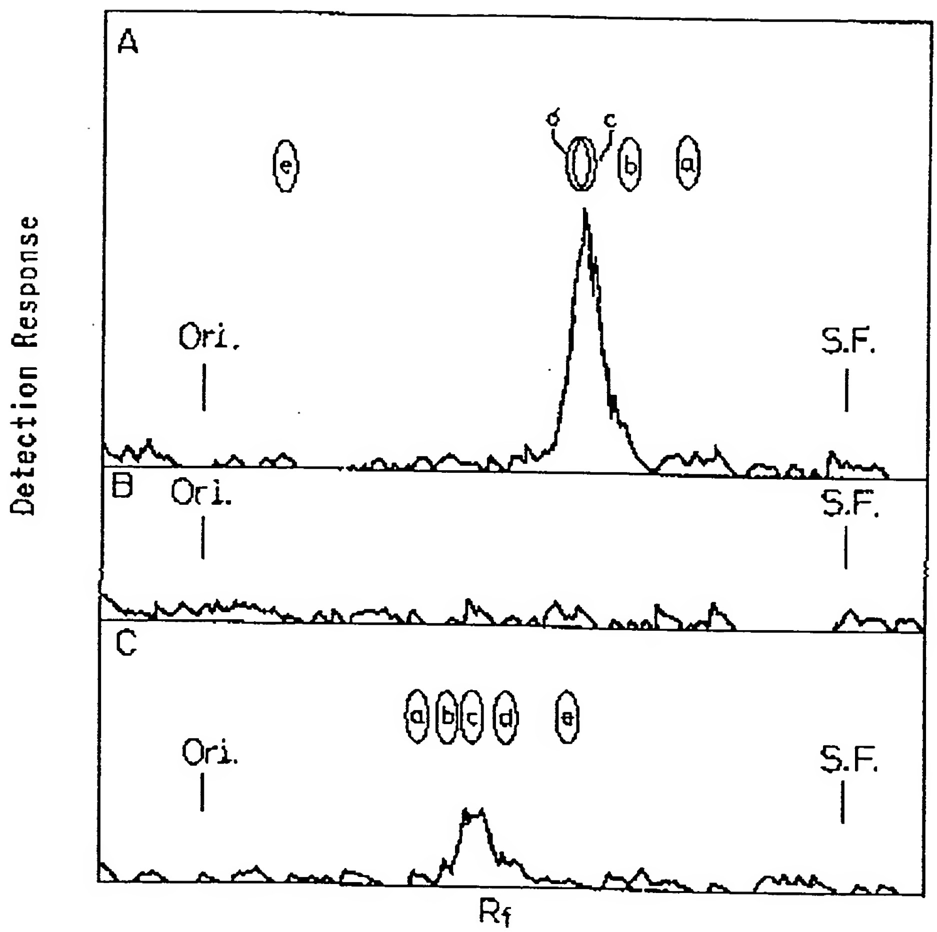


Fig.3

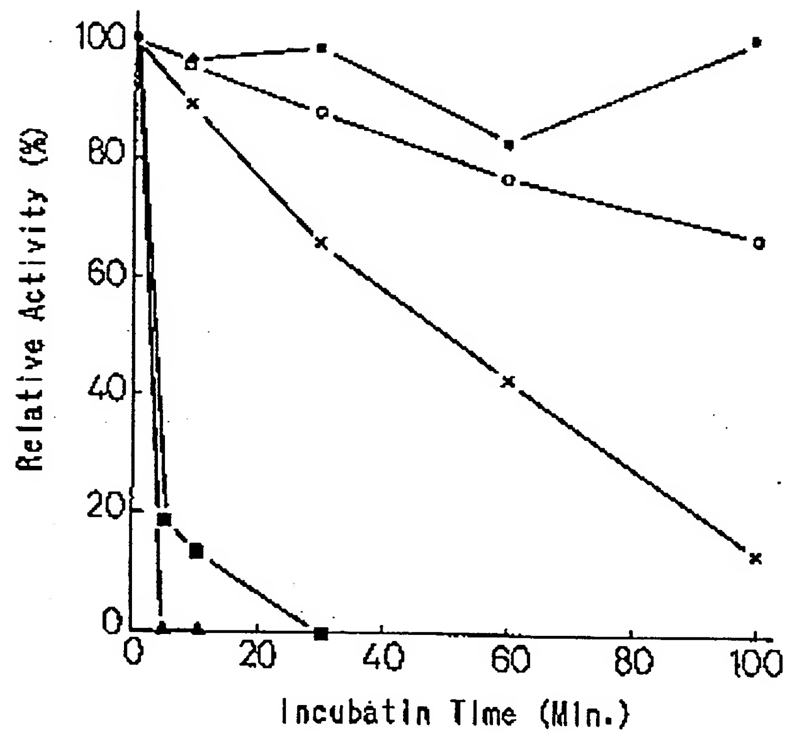

Relative Activity (%)
100
80
60
40
20
0
0
20
40
60
80
100
Incubatin Time (Min.)

Fig.4

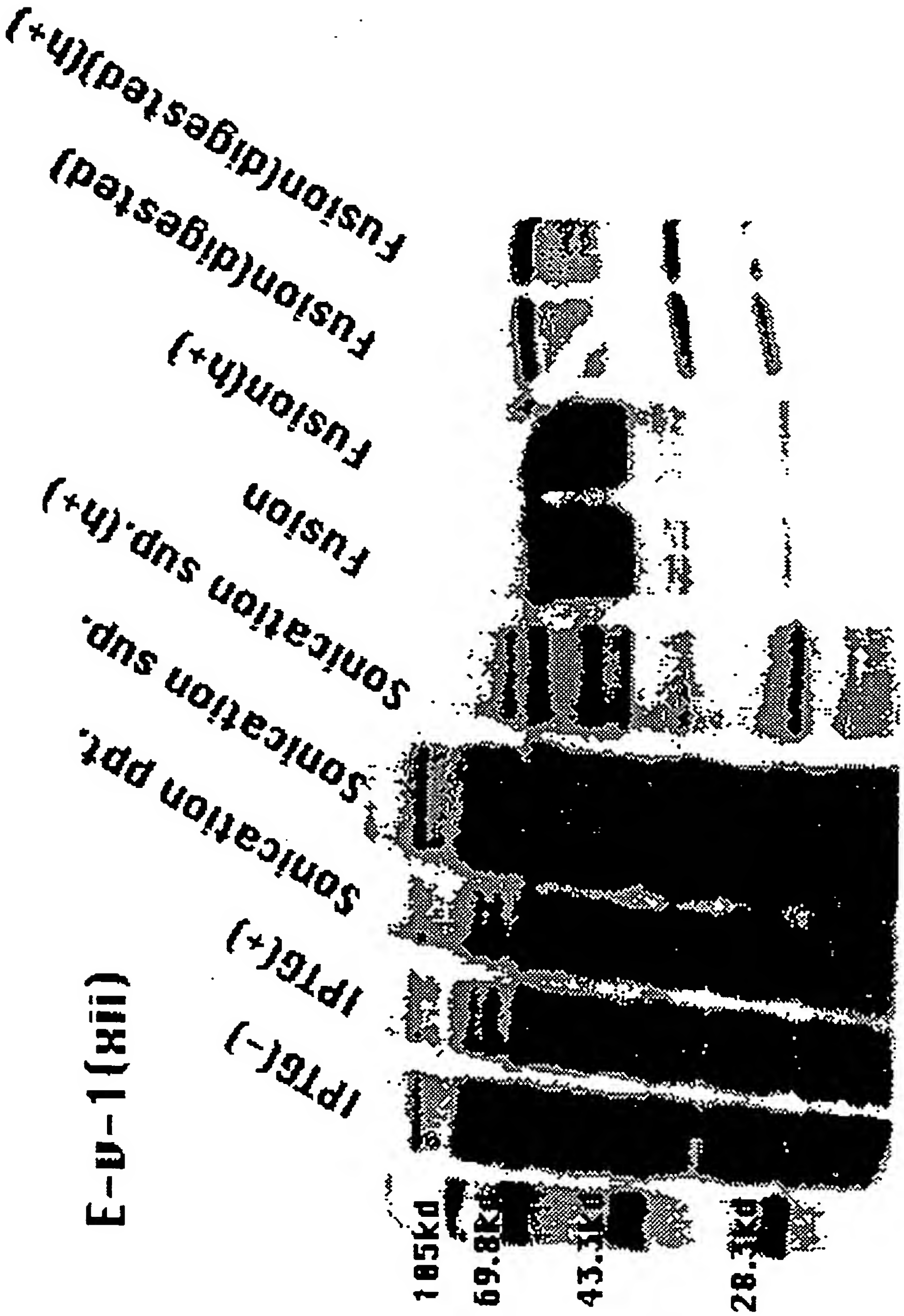

E-U-1(xiii)
IPTG(-)
IPTG(+)
Sonication ppt.
Sonication sup.
Sonication sup.(h+)
Fusion
Fusion(h+)
Fusion(digested)
Fusion(digested)(h+)
185kd

Fig.5

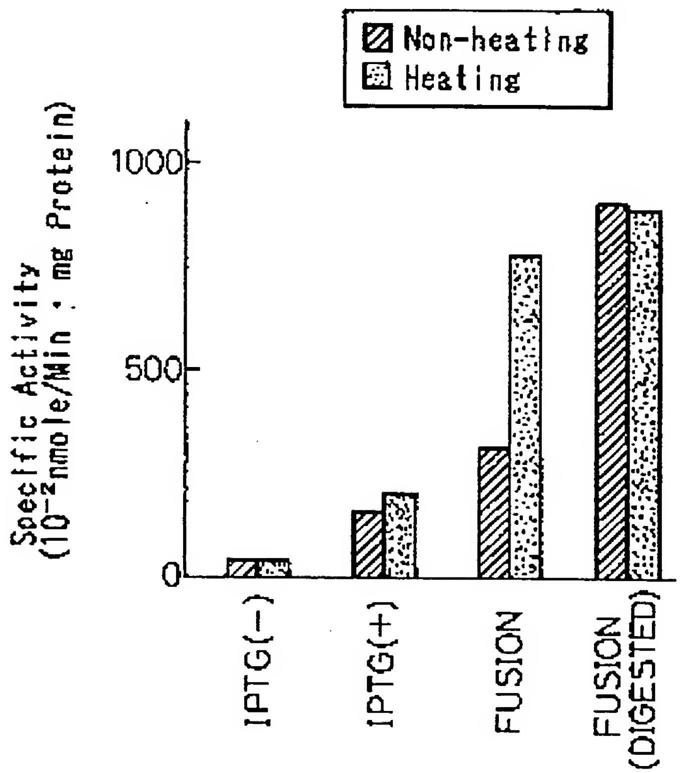

Non-heating
Heating
Specific Activity
(10^-2 nmole/Min · mg Protein)
1000
500
0
IPTG(−)
IPTG(+)
FUSION
FUSION
(DIGESTED)

Fig.6

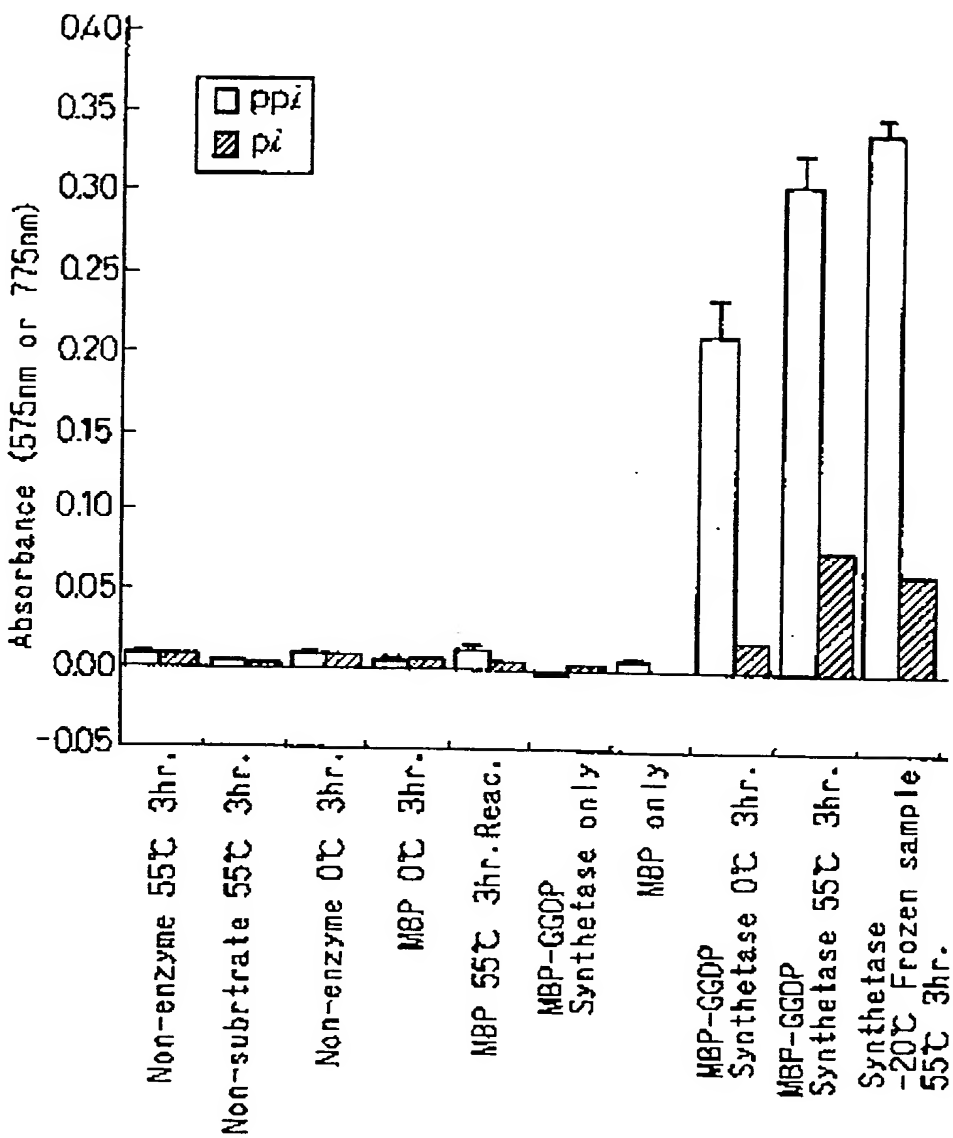

0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
-0.05
Absorbance (575nm or 775nm)
ppi
pi
Non-enzyme 55℃ 3hr.
Non-subrtrate 55℃ 3hr.
Non-enzyme 0℃ 3hr.
MBP 0℃ 3hr.
MBP 55℃ 3hr. Reac.
MBP-GGDP Synthetase only
MBP only
MBP-GGDP Synthetase 0℃ 3hr.
MBP-GGDP Synthetase 55℃ 3hr.
Synthetase -20℃ Frozen sample 55℃ 3hr.

Fig.7


F_OH
GG_OH
C(t)
D(t)
E(t)
C(d)
D(d)
E(d)
primer:GPP
C(t)
D(t)
E(t)
primer:(all-E)-FPP
C(d)
D(d)
E(d)
C(t)
D(t)
E(t)
C(d)
D(d)
E(d)
primer:(2Z,6E)-FPP